Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 119 (2017) 30–41

Procedia Computer Science

www.elsevier.com/locate/procedia

6th International Young Scientists Conference in HPC and Simulation, YSC 2017, 1-3 November 2017, Kotka, Finland

Multi-View Data approaches in Recommender Systems: an Overview

(Invited Paper)

Iván Palomares*[a], Sergey V. Kovalchuk[b]

[a] *School of Computer Science, Electrical and Electronic Engineering, and Engineering Mathematics. University of Bristol, Bristol, United Kingdom*

[b] *ITMO University, Saint-Petersburg, Russian Federation*

Abstract

This paper overviews an assortment of recent research work undertaken on recommender system models based on using multiple views of user and item-related data across the recommendation process. A summary of representative literature on multi-view recommender approaches is provided, describing their main characteristics, such as: their potential to overcome most common shortcomings in conventional recommender systems, as well as the use of data science, learning techniques and aggregation processes to combine information stemming from multiple views. A tabular summary is provided to facilitate the comparison of the similarities and differences among the surveyed works, along with commonly identified directions for future research in the topic.

Peer-review under responsibility of the scientific committee of the 6th International Young Scientist conference in HPC and Simulation

Keywords: Recommender Systems; Collaborative Filtering; Clustering; Multi-View Data; Multi-View Recommendation; User Similarity; User Trust; Aggregation Functions

1. Introduction

As the availability of digital information, resources and on-line content continuously increases, users have access to a wealth of information. The sheer volume and variety of content available however can make it difficult for them to find information that suitably meet their needs. In these circumstances, Recommender Systems (RS) arose to overcome such challenges, nowadays playing an important role in myriad e-commerce, personalization and decision-making domains [2, 9]. There exist a vast array of applications of RS, ranging from the most widely known scenarios

* Corresponding author. Tel.: +44 (0117) 331 5055
E-mail address: i.palomares@bristol.ac.uk

1877-0509 © 2018 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the 6th International Young Scientist conference in HPC and Simulation
10.1016/j.procs.2017.11.157

(recommending products, movies, music, etc.) to much more specialized domains, e.g. recommending best practices for urban resilience [32] and urban sustainable development initiatives.

As recommender, decision support and Web systems have progressed and improved in terms of sophistication and connectivity with other systems, the quantity and quality of feature data available to RS to make recommendations has also expanded and improved dramatically [23]. Moreover, the ever-increasing explosion of readily available information about users and items in the Internet make it more necessary than ever before to incorporate and combine multiple views or dimensions of such information (e.g. ratings, social trust, textual and multi-media information) in the processes typically undertaken by conventional recommender models [19, 20, 32]. This may not only improve recommendation accuracy and quality, but also might in some cases alleviate some of the most frequently found shortcomings and vulnerabilities in recommender approaches. Unsurprisingly, several scholars have recently focused their efforts on recommender domains in which multiple views of information shall be exploited meaningfully to produce more accurate recommendations amid diverse situations. Such approaches are in most cases referred to as *multi-view* RS methods [13, 19, 27, 29]. Whilst there is no shortage of literature surveys on major or more generic families of RS approaches, such as Collaborative Filtering (CF) or content-based [7, 10, 46], to our knowledge no theoretical work has been undertaken to date on specifically gathering and compiling a summary of representative research on multi-view RS models and methods.

This paper focuses on RS research based on the use of multi-view data approaches. In particular, we provide a concise overview of recent recommender system approaches characterized by integrating multiple views of user and item-related data at various stages of the recommendation process. The summary of related literature provided consists of 14 selected works handling multiple views of data. Aspects such as the management of common limitations and drawbacks in conventional recommender systems, the employment of data science and learning techniques for knowledge extraction, and the use of flexible aggregation strategies to combine information from multiple views, are particularly pointed out. A tabular comparison of similarities, conclusions in common and differences among the surveyed works is also presented, along with commonly identified directions for future research in the topic.

This paper is organized as follows: Section 2 reviews some basic preliminares on RS and aggregation operators. Section overviews the 14 selected works on multi-view based RS approaches, highlighting their most relevant characteristics on the collection, use and fusion of multiple views of data across the recommendation process. Finally, Section 4 concisely summarizes both common and differentiating aspects among the reviewed works and points out some directions for future research on multi-view RS.

2. Preliminaries

2.1. Basic Concepts on Recommender Systems

RSs attempt to filter items to users, by predicting a rating value for unseen items by such users so as to filter and rank the “best” unrated items in terms of their prediction value. Examples of existing RS techniques include, but are not limited to:

- *Content-based*: They recommend items that are similar to those positively rated by the user [26].
- *Collaborative filtering (CF) based*: They recommend items positively rated by similar users to the target user [15, 39]. CF approaches can be further classified into two subtypes [46]:
  - *Model-based CF*: These approaches use user-item rating information to learn a prediction model.
  - *Neighborhood-based*: The approaches use user-item ratings to directly predict ratings for unseen items, based on identifying the most similar users to the target user.
- *Knowledge-based*: They suggest items based on inference on the user needs and preferences [9].
- *Demographic*: They provide recommendations based on the demographic profile of users [41].
- *Context-aware*: They consider contextual information (location, time, etc.) in the recommendation process. Context-aware recommender systems are typically hybridized with other techniques, such as CF [2].
- *Clustering-based*: Commonly viewed as a variant of CF methods, clustering-based recommendation models create a overall similarity-based clustering of the user space (e.g. based on rating information), instead of determining the neighbors or most similar users to a target user [19].

Hybrid approaches and Group Recommender System models, such as collaborative filtering and knowledge-based, or collaborative filtering and demographic, have been subject of extensive research in recent years [12, 9]. Most conventional RS models typically consist of the following three sources of information:

- A set of *items*, $X = \{x_1, \dots x_l\}$ (e.g. products, services or other information resources), which may be defined by metadata or other type of information about them.
- A set of *users* of the system, $U = \{u_1, \dots, u_m\}$ who may provide information about themselves, both explicitly (e.g. age, gender, zip code), and implicitly (i.e. preferences over items).
- A set of users' preferences or *ratings* over the items, $R \subseteq U \times I \to D$, expressed as a value in a rating domain D, indicating the preference or satisfaction degree of a particular user with a specific item.

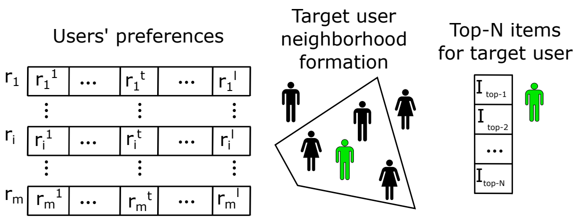

Fig. 1. Collaborative Filtering process in RS

Most multi-view approaches reviewed in this study are based on extending neighborhod-based CF approaches, hence CFRS are now reviewed in further detail. Neighborhood-based CFRS are based on similarity between users [1, 38]. These methods take the users' preferences over items or *rankings* as input for predicting (recommending) new items that might potentially be of interest to them, based on items positively rated by similar users or (*neighbors*). The underlying premise is that those items yet unknown to a target user $u_i \in U$, $i = 1, \dots, m$, and positively rated by similar users, might be foreseen as satisfactory to her/him. There exists an assortment of probabilistic and non-probabilistic approaches for CFRS, such as nearest neighbor-based models, dimensionality reduction models, Bayesian models, etc [22]. A common approach in CFRS is the k-nearest neighbor (*kNN*) collaborative filtering, also known as user-user collaborative filtering [8], which determines a neighborhood or subset of similar users to the target user. Central to the neighborhood process in kNN-based CFRS, is the use of an adequate *similarity measure* [15]. The Pearson correlation coefficient, cosine similarity and Spearman rank are well-known examples of similarity measures commonly utilized in related literature. For instance, the Pearson correlation coefficient among two users $u_i, u_j \in U$, whose subset of commonly rated items is denoted by $X_{i,j}$, is classically calculated as follows:

$$sim(u_i, u_j) = \frac{\sum_{x_t \in X_{i,j}} (r_i^t - \overline{r}_i)(r_j^t - \overline{r}_j)}{\sqrt{\sum_{x_t \in X_{i,j}} (r_i^t - \overline{r}_i)^2} \sqrt{\sum_{x_t \in X_{i,j}} (r_j^t - \overline{r}_j)^2}} \quad (1)$$

where r_i^t is the rating provided by u_i on item x_t, and $\overline{r}_i$ is the average rating expressed by u_i. Based on the set of neighbor users to u_i, a *prediction function* is utilized to predict a rating value for each item not rated yet by u_i. The most frequently utilized prediction function is a weighted sum function that aggregates k neighbor users' ratings over an item by using similarities as weights:

$$p(u_i, x_t) = \frac{\sum_{j=1}^{k} sim^{i,\sigma(j)} \cdot r_{\sigma(j)}^t}{\sum_{j=1}^{k} sim^{i,\sigma(j)}} \quad (2)$$

where $u_{\sigma(j)}$ denotes the j-th neighbor user in u_i's neighborhood. As a result, a list of recommendations is delivered by decreasing order of such a prediction value [39]. Figure 1 illustrates the operation scheme of a CFRS, according to which items are ranked for a target user based on neighborhood formation and user preference similarity.

Users' ratings in CFRSs can take different forms, depending on the system and application domain in which they are implemented [15]. In many domains, numerical ratings such as a 1-5 numerical scale are typically adopted. By contrast, *implicit or unary ratings*, are common in e-commerce deployments: an item $x_t \in X, t = 1, \ldots, l$, is either rated by u_i (e.g. marked as favorite, purchased in online shops), or not rated (unknown or non-specified preference over x_t).

The *cold-start* problem has been subject of a considerable deal of research within the area of recommender systems, particularly in CFRS. This problems arises when the amount of available ratings is relatively small and hence insufficient to effectively apply traditional CF techniques [3]. Two main variants of the cold-start problem have been distinguished: the *new item* cold-start problem, which occurs when a new item has been introduced in the system and not enough users rated it; and the *new user* cold-start problem, which takes place when a newly registered user has rated a small or null number of items, hence the system is unable to produce meaningful recommendations for her/him [37]. Likewise, different types of methods have been proposed to deal with the issue, such as: making use of additional information sources, improving the user similarity methods, and using hybrid RS methods [40].

Other important problems commonly found in RS models, and requiring special attention, are:

- *Sparsity*: This occurs when the amount of available items is exceedingly large, hence the amount of ratings provided by users on items (including the most experienced and/or active users in the system) is too small to make reliable recommendations.
- *Diversity*: Diversification (e.g. by recommending a proportion of "unusual" items to the user) is a crucial aspect to consider in some recommender domains in order to ehance user experience avoid overfitting. Nevertheless, it is usually a sensible practice to strike a balance among diversity and quality in recommendations [24].
- *Shilling attacks*: Also known as profile injection attacks, they consist in introducing overly biased ratings on specific items to degenerate the recommendation accuracy and/or cause reputational damage. The study of resilient RS models to counter these attacks has been extensively tackled in recent literature [18].

2.2. Aggregation Operators

The fusion of information is an essential element in intelligent and decision support systems [14]. RS are no exception in applying aggregation techniques (e.g. via a similarity-weighted average for predictions, see Eq. (2))) to avail of different sources of information to produce meaningful recommendations [4]. The purpose of aggregation functions is to combine a n-tuple of values or elements belonging to a set (e.g. unit interval [5]) into a single representative value.

[5] An aggregation function in the unit interval is a mapping $f : [0, 1]^n \rightarrow [0, 1]$, $n \geq 1$, producing an output value from a set of n input values $A = a_1, \ldots, a_n$. Every aggregation function in the [0,1] interval satisfies the following three properties:

(i) **Identity when Unary**: $f(a) = a$.
(ii) **Boundary**: $f(0, \ldots, 0) = 0$ and $f(1, \ldots, 1) = 1$.
(iii) **Monotonicity or Non-decreasing**: $a_z \leq b_z\ \forall z = 1, \ldots, n$ implies $f(a_1, \ldots, a_n) \leq f(b_1, \ldots, b_n)$.

Typically, aggregation in RS has been undertaken to combine similarity and rating information, by applying prototypical functions such as the arithmetic or weighted mean. However, in some contexts, particularly those in which information from multiple views or dimensions must be aggregated [32], it is desirable a function that fulfills additional mathematical properties, for instance:

1. **Idempotence**: f$(a, a, \ldots, a) = a$.
2. **Compensation**: $\min_z a_z \leq f(a_1, \ldots, a_n) \leq \max_z a_z$.
3. **Associativity**: $f(a, f(b, c)) = f(f(a, b), c)$.
4. **Reinforcement**: Tendency of multiple high (*resp.* low) values to reinforce each other, leading to an even higher (*resp.* lower) result.

For the interested reader, we refer to [5, 36] for a comprehensive overview of the main classes of aggregation functions.

Below we briefly revise two families of aggregation functions, OWA and uninorm operators, which have been utilized by Palomares et al. in [32] to combine pairwise user similarity information stemming from user preference and user profile views. The OWA (Ordered Weighted Averaging) operators were introduced by Yager in [42], and they constitute a widely used family of weighted aggregation operators in the literature, particularly in multi-criteria decision support and fuzzy decision making [35]. Let $A = \{a_1, \ldots, a_n\}$ ($a_z \in [0,1]$) be a set of n values to aggregate. A OWA operator is a mapping $OWA_W : [0,1]^n \to [0,1]$, with an associated weighting vector $W = [w_1 w_2 \ldots w_n]^\top$, such that $w_z \in [0,1]$, $\sum_z w_z = 1$ and,

$$OWA_W(a_1, \ldots, a_n) = \sum_{z=1}^{n} w_z b_z \tag{3}$$

where b_z is the z-th largest value in A. OWA operators are characterized by assigning a weight w_z to the z-th largest element in A, unlike classic weighted average operators, which assign weight w_z to the z-th element in the input set, a_z (i.e. without previously sorting inputs in decreasing order).

The behavior of OWA operators can be flexibly defined and classified based on their weighting vector W (either optimistic, pessimistic or neutral). To determine the attitudinal character of the specific operator utilised, a measure called *orness*, denoted by $orness(W)$ was also introduced in [42]:

$$orness(W) = \frac{1}{n-1} \sum_{z=1}^{n} (n-z) w_z \tag{4}$$

Optimistic (OR-like) OWA operators are those where $orness(W) > 0.5$, whereas pessimistic (AND-like) operators have $orness(W) < 0.5$ [43]. The higher $orness(W)$, the more importance is assigned to the highest values in A, therefore the closer the aggregated result is to $\max(a_1, \ldots, a_n)$. Conversely, the lower $orness(W)$, the more importance is given to the highest values in A, and the closer the output is to $\min(a_1, \ldots, a_n)$.

A central aspect for the definition of an OWA operator consists in the construction of the weighting vector W. Different approaches have been proposed in the literature to facilitate their computation, e.g. by using fuzzy linguistic quantifiers or from learning approaches [42, 44]. Some special cases of OWA operators are [16]:

- The *maximum* operator, with $orness(W) = 1$, $w_1 = 1$ and $w_z = 0, z \neq 1$.
- The *minimum* operator, with $orness(W) = 0$, $w_n = 1$ and $w_z = 0, z \neq n$.
- The *arithmetic mean*, with $orness(W) = 0.5$ and $w_z = 1/n \; \forall z$.

Uninorm aggregation operators were introduced by Yager and Rybalov in [45, 17] to provide a generalization of the t-norm and the t-conorm operators [5]. Unlike t-norms and t-conorms, whose neutral elements are 1 and 0 respectively, uninorms have a neutral element $g \in [0,1]$ lying anywhere in the unit interval. Whilst OWA operators allowed to define varying attitudes within an averaging behavior, uninorm aggregation operators present a varying behavior (namely conjunctive, disjunctive or averaging), depending on the input values being higher or lower than g. A uninorm is a mapping, $\mathcal{U} : [0,1]^2 \to [0,1]$, having the following properties for all $a, b, c, d \in [0,1]$:

i) *Commutativity*: $\mathcal{U}(a,b) = \mathcal{U}(b,a)$.
ii) *Monotonicity*: $\mathcal{U}(a,b) \geq \mathcal{U}(c,d)$ if $a \geq c$ and $b \geq d$.
iii) *Associativity*: $\mathcal{U}(a, \mathcal{U}(b,c)) = \mathcal{U}(\mathcal{U}(a,b), c)$.
iv) *Neutral element*: $\exists g \in [0,1] : \mathcal{U}(a,g) = a$.

Because of their associativity property, uninorm operators are typically defined for $n = 2$, and additional input values can be successively aggregated without affecting the aggregated result. The conjunctive, disjunctive or averaging behavior depends on input values a, b being greater or lower than g. This distinctive property is graphically illustrated in Figure 2.

A notable characteristic of uninorm operators is their *full reinforcement* property: given any $g \in [0,1]$, uninorms show an *upward reinforcement* when both input values are high (above g), making the aggregated value even higher (disjunctive behavior). Conversely, they show a *downward reinforcement* when aggregating low input values (below g), so that the aggregated value is even lower (conjunctive behavior).

The cross-ratio uninorm is a continuous uninorm in $[0,1]^2 \backslash \{(0,1),(1,0)\}$, with neutral element $g = 0.5$:

$$\mathcal{U}(a,b) = \begin{cases} 0 & \text{if } (a,b) \in \{(0,1),(1,0)\}, \\ \frac{ab}{ab + (1-a)(1-b)} & \text{otherwise.} \end{cases} \tag{5}$$

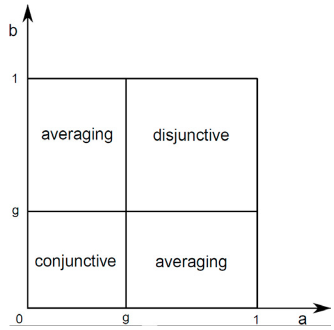

Fig. 2. Behavior of uninorm operators for different input values $a, b \in [0, 1]$ and a neutral element $g \in [0, 1]$

3. Recent Trends on Multi-View Data Fusion in RS

This section provides a concise literature review of multi-view RS research, focused on integrating multi-view data approaches as part of the personalization and recommendation process.

Oufaidia and Nouali proposed in [31] one of the earliest multi-view recommendation models. In order to overcome the cold start and sparsity problems that frequently undermine CFRS models, they presented a multi-view engine that, by exploiting semantic web technologies, incorporates three views of recommendation information: collaborative, social and semantic. Ontologies, tagging and social networks are some of the semantic web resources considered by the authors in their method. The collaborative (explicit or implicit ratings), socio-demographic and semantic data views constituting each user profile, are analyzed separately and independently to obtain *resp.* three user neighborhood models: collaborative neighborhood, social neighborhood and semantic neighborhood. Each of the three views produce their own recommendation lists, hence a ranking aggregation strategy is applied to obtain a hybrid, overall ranking of recommended items for a given user. Three possible strategies (mixed, weighted and switched) are proposed to do this, inspired by aggregation functions with different optimistic/pessimistic attitudes to obtain an overall ranking positions for those items which are recommended by multiple views simultaneously. The *BookCrossing* dataset, which contains over 42K instances of implicit rating data, is used to evaluate the model in conjunction with socio-demographic and semantic data generated synthetically. The precision is proved to improve with the proposed method, whereas recall is only improved when using a mixed hybridization strategy specifically.

Semantic data is also considered for recommendation processes in the work of Domingues et al. in [13]. In particular unstructured textual information pertaining items is mined to extract a topic item hierarchy, based on unsupervised learning. Two separate text clustering models are applied to obtain two co-association matrices, each of which represents a technical view (bag-of-words) and a privileged view (named entities), respectively. Both matrices are linearly combined at matrix element level, to obtain a single co-association matrix (describing a so-called “consensus clustering” of items). This is in turn utilized as a representation of relationships between documents that reflects both technical and privileged textual data views. Feature selection is subsequently used to derive a topic hierarchy, which is used as the similarity driven forcen for producing recommendations. A comparative evaluation against several baseline approaches is provided [13].

The (sometimes overlooked) goal of improving recommendation diversity is tackled by Li and Murata in [25]. In their study, the authors propose incorporating multi-dimensional clustering [11] into a CF model, so as to find a trade-off among accuracy and diversity in recommendations whilst enabling improvements in the latter. Multi-dimensional clustering approaches, such as subspace clustering, allows an object (e.g. user) to belong to multiple clusters across distinct subspaces. Predicated on the *Movielens* database for movie recommendation, it is illustrated that when user profile and item data have a large number of attributes, different clusterings can be generated according to different subsets of attributes. This idea has been illustrated by Li and Murata in Figure 3, based on which the overall recommendation process is divided into three phases: (1) *preprocessing and clustering* background data in the form of partitioned user and item profile data; (2) *cluster optimization*, with the removal of poor-quality clusters; (3) *collaborative filtering*, in which the target user’s preferences are analyzed and the attribute subspace (clustering dimension) most relevant to her/him is selected, after which a classical neighborhood and prediction approach is applied under that subspace.

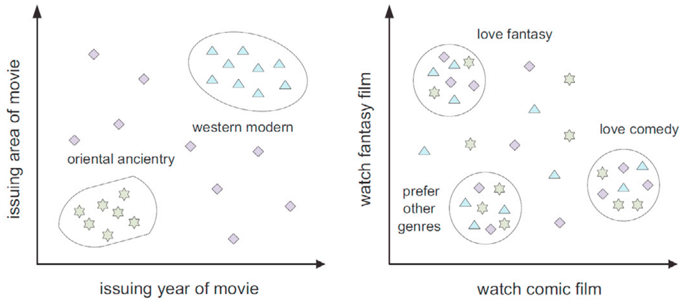

Fig. 3. Illustrating two user preference concepts via multi-dimensional clustering (taken from Li and Murata [25])

In [34], Qu et al. presented a multi-view semi-supervised recommender approach that exploits multimedia content in movies (image, text and audio), deeming each media type as a separate view space for movie recommendation. Once each view v_k is extracted from movie data, the similarity $sim_k(x_t, x_u)$ among items t and v under that view, is calculated, where sim is a similarity measure specific to the nature of the data in each view, for instance cosine similarity for the textual view. After having predicted a rating value $\tilde{r}^t_{i,k}$ on an item x_t by the ith user for each view v_k separately, predicted ratings across views are fused to obtain the overall prediction value $\tilde{r}^t_i$. An experimental study demonstrates that more comprehensive user profiles can be constructed by e.g. identifying their musical and visual preferences, which in turn enriches the (sometimes scarce) rating and explicit profile data associated with "cold" users.

Guo et al. in [19] investigated the problem of incorporating social relationship information in clustering-based methods for CF. They developed a multiview, clustering-based recommender approach that makes use of two dimensions of user information: (i) rating patterns, and (ii) social trust relationships. The *k-medoids* partitional clustering algorithm is applied to iteratively generate two different clusterings of them (one for each view), and then the resulting clusters from both views are combined through merging and pruning operations. It is noteworthy that in most contexts, trust values among users are binary in nature (trust links), hence Guo et al. define in their work the trust among users e_i, e_j based on their distance across the trust network. The predicted rating of an item x_t for a user u_i is calculated by an extension of the popular weighted sum prediction function, in which the importance weight $w_{i,j}$ of each neighbor user of u_i, $u_j \in C_i$, must be calculated as a combination of the rating similarity $s_{i,j}$ and trust degree $t_{i,j}$ among both users. Under the premise that a high weight $w_{i,j}$ requires both high similarity and high trust, the harmonic mean is used to aggregate both weights during prediction. In cases when a user appears in two clusters simultaneously, two different prediction values may arise on the same user u_i and item s_j. A regression problem is formulated to optimally combine such prediction values whilst minimizing the deviation from the actual user's preference on x_t [19]. Finally, to deal with the "new user" cold-start problem, Guo et al. propose a probabilistic approach that identifies the likelihood of a user belonging to a specific cluster predicating on preferential and trust data. The outperformance of the multiview clustering-based recommendation method is demonstrated in terms of accuracy and coverage.

A number of follow-up works incorporating trust information have been recently presented. In [20], Guo et al. introduced *TrustSVD*, the first extension of the state-of-the-art recommendation algorithm SVD++, incorporating social trust information. SVD++ also uses explicit and implicit rating information. Their study includes an empirical analysis that demonstrates the potential of trust and rating data to complement each other in a recommendation domain. Trust relationships between users - which are not necessarily symmetric/bidirectional- are exploited in the process of predicting an item rating for either a truster user or a trustee user, as illustrated in Figure 4. Experiments investigated by the authors in [20] utilize four different public datasets for movie recommendation. They also include the performance analysis of trust-based models in relation to the number of trusted neighbors per target user. Meanwhile, in [21] the authors focus the use of a trust-aware method on the ranking-based or *top-N item recommendation* problem (rather than the *rating prediction* problem), i.e. recommending an ordered list of relevant items to the user in question. Three factored similarity models are introduced based on matrix factorization techniques. A crucial hypothesis in [21] is that a user's social trust relationships play an influential role in the ranking score for an item. Their work includes an exhaustive comparative study, in which a total of 11 top-N item recommendation methods are compared against their proposed factored similarity models, concluding that item similarities and trust influence should receive more attention compared to user similarity, in order to achieve optimal performance in ranking-based frameworks.

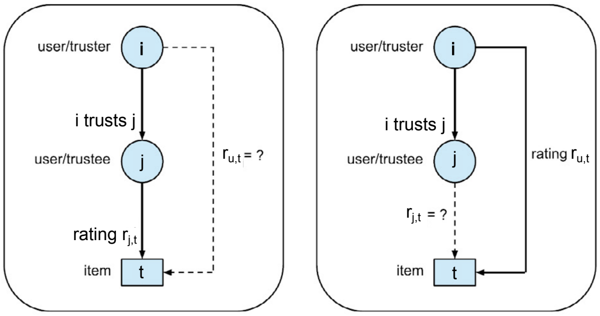

Fig. 4. Using trust relationships to help predicting ratings (adapted from [20])

The integration of trust information in CFRS models was also investigated by Moradi and Ahmadian in [30]. Several well-known shortcomings in traditional CF methods, including the sparsity problem and the existence of shilling attacks, are approached by introducing reliability measures on recommendations in such methods, predicated on similarity and trust statements. The proposed method firstly constructs a trust network, based on which initial ratings on unseen items are derived for a target user. The reliability measure is then used to determine the quality of predicted ratings, which in turn is utilized to reconstruct the trust network with the aim of further improving the recommendation accuracy, similarly to a feedback mechanism. A final rating prediction step is eventually applied to provide the user with top-*N* recommendations. The proposed reliability-based CF method is proved to outperform other similar approaches, both CF-based and trust-aware, whereas performance in coverage is reduced.

In [47], Zhang and Wang focused their efforts on alleviating the sparsity problem based on learning from multi-view data. They introduce a CF multi-view framework that combines the advantages of matrix factorization models and the abilities of transfer learning [33]. One of the strengths of their method relies in its ability to perform well in highly sparse rating contexts. The idea of transfer learning is to propagate "a priori" knowledge extracted from other related recommender systems into the target system, thus bridging information gaps across different systems. The information obtained by transfer learning processes is translated into multiple views of user-item rating matrices. As a result, embedding transfer learnign allows to automatically learn a multi-view model without the need for multiple views of data readily available in advance. The performance of Zhang and Wang's method is tested against the *CiteULike* and *LastFM* datasets, showing the improvements under the presence of multi-view content information exclusively, i.e. witout relying on other external sources such as social network data.

Berkani in [6] focused on CF techniques based on semantic and social dimensions. The author's approach relies in calculating the similarity between an user and his/her friends, such that the notion of friend has a twofold meaning (views): (1) individuals with domains of expertise and interests in common, and (2) other users with whom he/she has a strong degree of trust. Based on this, a total of three views are considered in their CF-based method:

- *Collaborative*: Based on calculating neighbors users with similar rating history, under an user-user memory-based CF approach.
- *Semantic*: It considers "friend" users with common interests and knowledge domains of expertise.
- *Social*: It identifies the "trusted friends" of a target user.

Thus, a collaborative filtering, semantic filtering and social filtering processes are undertaken in parallel to determine three separate neighborhoods. By assigning an importance degree to each of the three dimensions, the three neighborhoods or lists or recommended users are fused into an overall neighborhood. The work focuses exclusively on the problem of recommending like-minded users, whilst the intuitively subsequent process of predicting ratings on unseen items is not addressed.

Lu et al. consider the sheer amount of social data readily available nowadays upon the rapid development of microblogging systems. Accordingly, and since few works have focused on integrating microblogging data into recommender domains, they investigate in [27] multi-view user preference learning processes for *social recommendation* using microblogging data, to enhance recommender performance. In their work, multi-view refers to various descriptions of user preference, namely a *social view* (from microblogging systems data) and *recommend view* (from a product review site). User preferences in the social view are deemed as a low-dimensional representation of tagged information, whereas user preference in the recommend view are regarded as a low-dimensional representation of rating data. Both views have been previously evidenced as being strongly correlated, e.g. a user tagged as 'Geek' might be potentially interested in technology [27]. Two matrices are learnt from both views: a user-rating matrix enriched with side item information (from the recommender system), and a user-tag matrix (from the microblogging system). Both matrices are combined along with a third one, namely a user-user laplacian matrix representing social relationships, to finally obtain an aggregated user preference matrix, which can herein be used for e.g. neighborhood-based CF. Experiments are conducted using *Douban* (a chinese movie and music review site) and *Sina Weibo* (a microblogging system in China), both of which have a considerable number of registered users in common. Comparison with other baseline approaches demonstrate the computational learning efficiency of Lu et al.'s approach.

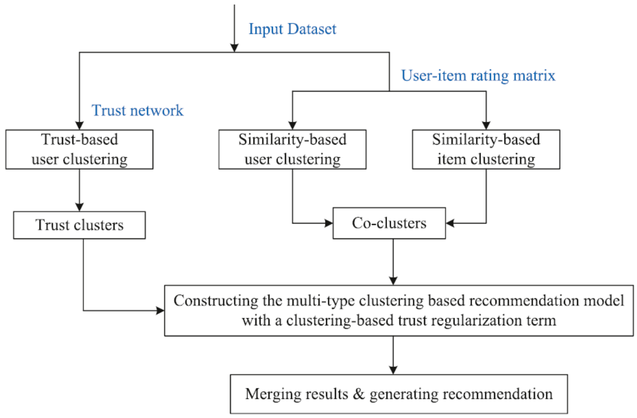

Fig. 5. A multi-type clustering-based recommender framework (taken from Ma et al. [29])

Ma et al. focus their research on clustering-based multi-view recommender methods, to tackle not only common drawbacks in CF methods, but also the relatively low accuracy that existing clustering-based methods still suffer to date. In [29] they developed a multi-type clustering-based unified recommender framework, that conflates similarity-based user clustering, similarity-based item clustering and trust-based user clustering. Contrary to traditional cluster methods, their multi-type clustering approach alleviates both the scarcity and cold start problems. The main two pieces of input are the user-item rating matrix and the social trust network. On the one hand, the user-item rating matrix is used to obtain two clusterings (similarity-based user clustering and similarity-based item clustering), which are subsequently combined into so-called co-clusters. On the other hand, trust-based user clusters are discovered via a SVD-based mining technique, which are in turn integrated with the previously obtained co-clusters into a multi-type clustering based recommendation model. In Figure 5, Ma et al. illustrated this process. The complexity and applicability of their model is demonstrated using consumer review website data. Moreover, in [28] the authors recently studied the explicit integration of both trust and distrust information to further improve clustering-based recommender models, arguing for instance that distrust relationship can be inferred from pairs of users allocated in different clusters. It is also illustrated how sparse rating matrices can be further completed by aggregating trust information pertaining trust neighborhoods among users.

Palomares et al. presented in [32] a multi-view CF model that calculates pairwise user similarity based on two data sources: users preferences (unary ratings) and user profile. User profile data consist of a finite number of information fields, such that the similarity among two users is calculated for each of these fields and then aggregated using an instance of OWA operator. Subsequently, the overall profile similarity is combined with the preference similarity by using a uninorm operator, which reinforces upwards (resp. downwards) the aggregated similarity in the cases when both views of user similarity are high (resp. low). The resulting framework is integrated with a Web platform for urban resilience resource recommendation. The multi-view similarity aggregation process is illustrated in Figure 6.

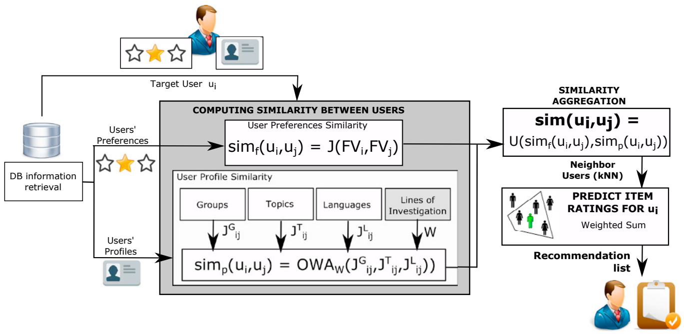

Fig. 6. Multi-view CF framework based on aggregation of similarity views [32]

Table 1 summarizes the most notable characteristics of the multi-view RS methods reviewed in this paper.

Table 1. Summary of reviewed multi-view RS literature

	Base approach	Data Views	Fusion of Views	Target problem(s) tackled	Other technique(s)
Oufaida and Nouali [31]	CF	Collaborative, Social, Semantic	Combine recommendation lists	Cold start, sparsity	-
Domingues et al. [13]	CB	Textual: technical (bag-of-words) and privileged (named entities)	Linear combination of co-association matrices	-	text clustering
Li and Murata [25]	CF, Clustering	Subsets of attributes (subspaces)	Choice of most relevant subspace to target user	Diversity	-
Qu et al. [34]	CB	Media types (audio, text, image)	Aggregation of predicted ratings	"New user" cold-start	-
Guo et al. [19]	CF, Clustering	Ratings, social trust	Merge clusterings, aggregate importance weights for predictions	"New user" cold-start	Regression for parameter optimization
Guo et al. [20, 21]	CF, Clustering	Explicit and Implicit Ratings, social trust	Merge clusterings, aggregate importance weights for predictions	-	Single Value Decomposition (SVD), Matrix Factorization
Moradi and Ahmadian [30]	CF	Ratings, social trust	N/A (mutual feedback among rating and trust views)	Sparsity, Shilling	-
Zhang and Wang [47]	CF	User-item ratings from multiple systems	N/A (transfer learning across views/systems)	Sparsity	Matrix factorization
Berkani [6]	CF	Collaborative, Semantic, Social	Weighted aggregation of neighborhoods	-	-
Lu et al. [27]	CB, CF	Social (microblogging system), recommend (reviews site)	Fusing user data matrices	-	-
Ma et al. [29]	CF, Clustering	User similarity, item similarity, social trust	Fusing clusterings	Sparsity, cold start	SVD
Ma et al. [28]	Clustering	Social trust and distrust	Fusing clusterings, aggregation of trust information	Sparsity	-
Palomares et al. [32]	CF	Ratings, User profile	Aggregating user similarities across views	-	OWA, Uninorm Operators

4. Discussion and Concluding Remarks

Based on the summary provided in Table 1, the following conclusions are drawn regarding both aspects in common and differences among the overviewed works.

- Clearly, the objective of improving recommender performance under the presence of common RS weaknesses and vulnerabilities (particularly the cold-start and sparsity problems) are a major motivation behind most of the surveyed multi-view approaches.
- Depending on the specific method, multi-view data (or the information derived from them) can be fused or unified not only by using a variety of techniques, but also at very diverse stages across the overall recommendation process, e.g. fusing recommendation lists after calculating predictions in [31], aggregating predicted ratings to obtain a single recommendation list in [34], merging multiple clustering results into one in [19], aggregating pairwise user similarity degrees from multiple views in [32], etc.
- Incorporating a social view, particularly related to social trust data, is a common feature found across several works, most of which also focus on extending classical CF approaches [6, 19, 20, 21, 28, 30, 31].
- The integration of clustering-based recommendation processes is another notable feature found in several multi-view works [19, 20, 21, 25, 28, 29].

- Interestingly, data science and machine learning methods play an important role in several of the reviewed multi-view approaches, not only limited to techniques frequently used in RS (SVD, matrix factorization), but also incorporating other techniques, such as transfer learning [47], text clustering [13] and regression of parameter values [19].

Furthermore, we conclude the paper pointing out some directions of research deserving further attention for the improvement of multi-view RS approaches.

1. Aggregation operators [4, 5] have proved in [32] to meaningfully reflect different aggregation attitudes in the process of recommending items to users, based on principles frequently applied in multi-criteria decision making. Thus, further exploring the ample catalogue of aggregation operators in distinct recommender domains poses an interesting direction of research in multi-view RS.
2. Decision support applications such as urban sustainable development, and user personalization in IoT and SmartCities environments.
3. Digital health applications, e.g. recommending personalized activity and health plans to users through wearable technologies, based on preferential data, activity trends and vital signs.
4. Extensions of multi-view approaches to group recommender systems, in which recommendation lists are jointly provided for a collective of users.

References

[1] Adomavicius, G., Kwon, Y., 2007. New recommendation techniques for multicriteria rating systems. IEEE Intelligent Systems 22, 48–55.
[2] Adomavicius, G., Tuzhilin, A., 2011. Context-Aware Recommender Systems. In F. Ricci (Ed.): Recommender Systems Handbook. Springer US. pp. 217–253.
[3] Aggarwal, C., 2016. Recommender Systems: The Textbook. Springer.
[4] Beliakov, G., Calvo, T., James, S., 2011. Aggregation of preferences in recommender systems, in: Recommender Systems Handbook, Springer.
[5] Beliakov, G., Pradera, A., Calvo, T., 2007. Aggregation Functions: A Guide for Practitioners. Springer.
[6] Berkani, L., 2015. SSCF: A semantic and social-based collaborative filtering approach, in: 2015 IEEE/ACS 12th International Conference of Computer Systems and Applications (AICCSA), pp. 1–4.
[7] Bobadilla, J., et al., 2013. Recommender systems survey. Knowledge-based Systems 46, 109–132.
[8] Breese, J.S., Heckerman, D., Kadie, C., 1998. Empirical analysis of predictive algorithms for collaborative filtering, in: Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence, Morgan Kaufmann Publishers Inc.. pp. 43–52.
[9] Burke, R., 2007. Hybrid Web Recommender Systems. Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 377–408.
[10] Chen, H., Parra, D., Verbert, K., 2016. Interactive recommender systems: a survey of the state of the art and future research challenges and opportunities. Expert Systems with Applications 56, 9–27.
[11] Chen, T., Zhang, N.L., Liu, T., Poon, K.M., Wang, Y., 2012. Model-based multidimensional clustering of categorical data. Artificial Intelligence 176, 2246 – 2269.
[12] Domingues, M.A., Manzato, M.G., Marcacini, R.M., Sundermann, C.V., Rezende, S.O., 2014. Using contextual information from topic hierarchies to improve context-aware recommender systems, in: 2014 22nd International Conference on Pattern Recognition (ICPR), IEEE. pp. 3606–3611.
[13] Domingues, M.A., Sundermann, C.V., Barros, F.M.M., Manzato, M.G., Pimentel, M.G.C., Rezende, S.O., Oliveira, S., 2015. Applying multi-view based metadata in personalized ranking for recommender systems, in: Proceedings of the 30th Annual ACM Symposium on Applied Computing, pp. 1105–1107.
[14] Doumpos, M., Grigoroudis, E., 2013. Multicriteria Decision Aid and Artificial Intelligence. Wiley.
[15] Ekstrand, M., Riedl, J., Konstan, J., 2010. Collaborative filtering recommender systems. Human-Computer Interaction 4, 81–173.
[16] Fodor, J., Marichal, J.L., Roubens, M., 1995. Characterization of the ordered weighted averaging operators. IEEE Transactions on Fuzzy Systems 3, 236–240.
[17] Fodor, J., Yager, R., Rybalov, A., 1997. Structure of uninorms. International Journal of Uncertainty, Fuzziness and Knowledge-based Systems 5, 411–427.
[18] Gunes, I., Kaleli, C., Bilge, A., Polat, H., 2014. Shilling attacks against recommender systems: a comprehensive survey. Artificial Intelligence Review 42, 767–799.
[19] Guo, G., Zhang, J., Yorke-Smith, N., 2015. Leveraging multiviews of trust and similarity to enhance clustering-based recommender systems. Knowledge-based Systems 74, 14–27.
[20] Guo, G., Zhang, J., Yorke-Smith, N., 2016. A novel recommendation model regularized with user trust and item ratings. IEEE Transactions on Knowledge and Data Engineering 28, 1607–1620.
[21] Guo, G., Zhang, J., Zhu, F., Wang, X., 2017. Factored similarity models with social trust for top-n item recommendation. Knowledge-Based Systems 122, 17 – 25.

[22] Herlocker, J.L., Konstan, J.A., Riedl, J., 2000. Explaining collaborative filtering recommendations, in: Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work, pp. 241–250.
[23] Hu, Y., Peng, Q., Hu, X., Yang, R., 2015. Time aware and data sparsity tolerant web service recommendation based on improved collaborative filtering. IEEE Transactions on Services Computing 8, 782–794.
[24] Kunaver, M., Porl, T., 2017. Diversity in recommender systems a survey. Knowledge-Based Systems 123, 154 – 162.
[25] Li, X., Murata, T., 2012. Using multidimensional clustering based collaborative filtering approach improving recommendation diversity, in: 2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, pp. 169–174.
[26] Lops, P., De Gemmis, M., Semeraro, G., 2011. Content-based recommender systems: State of the art and trends, in: Recommender systems handbook. Springer, pp. 73–105.
[27] Lu, H., Chen, C., Kong, M., Zhang, H., Zhao, Z., 2016. Social recommendation via multi-view user preference learning. Neurocomputing 216, 61 – 71.
[28] Ma, X., Lu, H., Gan, Z., Zeng, J., 2017. An explicit trust and distrust clustering based collaborative filtering recommendation approach. Electronic Commerce Research and Applications 25, 29 – 39.
[29] Ma, X., Lu, H., Gan, Z., Zhao, Q., 2016. An exploration of improving prediction accuracy by constructing a multi-type clustering based recommendation framework. Neurocomputing 191, 388 – 397.
[30] Moradi, P., Ahmadian, S., 2015. A reliability-based recommendation method to improve trust-aware recommender systems. Expert Systems with Applications 42, 7386 – 7398.
[31] Oufaida, H., Nouali, O., 2009. Exploiting semantic web technologies for recommender systems: A multi view recommendation engine, in: Proceedings of the 7th International Conference on Intelligent Techniques for Web Personalization & Recommender Systems - Volume 528, pp. 87–92.
[32] Palomares, I., Browne, F., Wang, H., Davis, P., 2015. A collaborative filtering recommender system model using owa and uninorm aggregation operators, in: Intelligent Systems and Knowledge Engineering (ISKE), 2015 10th International Conference on, pp. 382–388.
[33] Pan, S., Yang, Q., 2009. A survey on transfer learning. IEEE Transactions on Knowledge and Data Engineering 22, 1345–1359.
[34] Qu, W., Song, K.S., Zhang, Y.F., Feng, S., Wang, D.L., Yu, G., 2013. A novel approach based on multi-view content analysis and semi-supervised enrichment for movie recommendation. Journal of Computer Science and Technology 28, 776–787.
[35] Roubens, M., 1997. Fuzzy sets and decision analysis. Fuzzy Sets and Systems 90, 199–206.
[36] Rudas, I.J., Pap, E., Fodor, J., 2013. Information aggregation in intelligent systems: An application oriented approach. Knowledge-Based Systems 38, 3 – 13. Special Issue on Advances in Fuzzy Knowledge Systems: Theory and Application.
[37] Safoury, L., Salah, A., 2013. Exploiting user demographic attributes for solving cold-start problem in recommender system. Lectures Notes in Software Engineering 1, 303–307.
[38] Sarwar, B., Karypis, G., Konstan, J., Riedl, J., 2001. Item-based collaborative filtering recommendation algorithms, in: Proceedings of the 10th international conference on World Wide Web, ACM. pp. 285–295.
[39] Schafer, J.B., Frankowski, D., Herlocker, J., Sen, S., 2007. Collaborative filtering recommender systems, in: The adaptive web. Springer, pp. 291–324.
[40] Son, L.H., 2016. Dealing with the new user cold-start problem in recommender systems: A comparative review. Information Systems 58, 87 – 104.
[41] Vozalis, M.G., Margaritis, K.G., 2007. Using svd and demographic data for the enhancement of generalized collaborative filtering. Information Sciences 177, 3017–3037.
[42] Yager, R., 1988. On ordered weighted averaging aggregation operators in multi-criteria decision making. IEEE Transactions on Systems, Man and Cybernetics 18, 183–190.
[43] Yager, R., 1993. Families of OWA operators. Fuzzy Sets and Systems 59, 125–148.
[44] Yager, R., 1996. Quantifier guided aggregation using OWA operators. International Journal of Intelligent Systems 11, 49–73.
[45] Yager, R., Rybalov, A., 1996. Uninorm aggregation operators. Fuzzy Sets and Systems 80, 111–120.
[46] Yang, X., et al., 2014. A survey of collaborative filtering based social recommender systems. Computer Communications 41, 1–10.
[47] Zhang, Q., Wang, H., 2015. Collaborative Multi-view Learning with Active Discriminative Prior for Recommendation. Springer International Publishing. pp. 355–368.